
MACHINE LEARNING DIFFERENCE CHARGE DENSITY

Xiwen Li¹, LiangLiang Hong¹, Yingwei Chen¹, Hongjun Xiang^{1,*}

¹ Key Laboratory of Computational Physical Sciences (Ministry of Education), Institute of Computational Physical Sciences, State Key Laboratory of Surface Physics, and Department of Physics, Fudan University, Shanghai 200433, China

* E-mail: hxiang@fudan.edu.cn

ABSTRACT

In density functional theory(DFT), the ground state charge density is the fundamental variable which determines all other ground state properties. Many machine learning charge density models are developed by prior efforts, which have been proven useful to accelerate DFT calculations. Yet they all use the total charge density (TCD) as the training target. In this work, we advocate predicting difference charge density (DCD) instead. We term this simple technique by Δ -SAED, which leverages the prior physical information of superposition of atomic electron densities (SAED). The robustness of Δ -SAED is demonstrated through evaluations over diverse benchmark datasets, showing an extra accuracy gain for more than 90% structures in the test sets. Using a Si allotropy dataset, Δ -SAED is demonstrated to advance model’s transferability to chemical accuracy for non-self-consistent calculations. By incorporating physical priors to compensate for the limited expressive power of machine learning models, Δ -SAED offers a cost-free yet robust approach to improving charge density prediction and enhancing non-self-consistent performance.

1 Introduction

In the past decade, machine learning (ML) methods have gained popularity in materials modeling due to their favorable $O(N)$ scaling in supervised tasks [1], thanks to the so-called nearsightedness principle of electronic matter [2]. A major source of training data for ML methods comes from density functional theory(DFT) calculations. DFT is widely used in quantum atomistic simulations owing to its good balance between accuracy and efficiency [3]. However, the $O(N^3)$ scaling for solving the Kohn-Sham equations [4] self-consistently hindered its application to large scale simulations. To address this, various ML approaches have been developed to learn the mapping from atomic structures to DFT-computed properties directly. Among these, neural network potentials (NNPs), trained to predict formation energies and atomic forces, are the most extensively studied and form the foundation of other ML methods in materials modeling [5, 6]. The success of NNPs has inspired researchers to adapt their architectures to learn other properties, such as the ground state charge density [7].

Accurate ML charge density is desired for several reasons. Firstly, the first Hohenberg-Kohn theorem states that the ground-state charge density uniquely determines all ground-state properties of an interacting electron system [8, 9]. In principle ML charge density could allow us to bypass the self-consistent procedure in Kohn–Sham formalism to obtain ground-state properties by one diagonalization process. However, DFT calculations are extremely sensitive to the self-consistent charge density. In practice, the more accurate the ML-predicted charge density, the more likely it is to achieve self-consistency, and the more reliable the properties are derived from non-self-consistent calculations. Secondly, the ground-state charge density can be used further as descriptors for downstream machine learning tasks. In NNPs this indirect approach has been reported to outperform significantly [7, 10, 11]. Lastly, encoding rich electronic structure and chemical bonding information, charge density—along with its derived quantities such as Bader charge [12], electron localization function [13] and Hartree potential—can be used directly to characterize a wide range of phenomena including polarization [14], defect states [15], charge transfer in photovoltaics [16] and lithium batteries [17], to name a few.

Chronologically, the architecture of ML charge density models has evolved from kernel regression to message passing neural networks, and from obeying invariance to incorporating equivariant hidden layers, for better transferability. Meanwhile, the choice of fitting targets ranges from basis expansion coefficients to grid representations, depending on the trade-off between efficiency and accuracy. Early triumphs in modeling charge density use kernel ridge regression to fit basis expansion coefficients of charge density [7, 18–20]. By comparing unknown structures to those in the training set through a kernel function, kernel-based ML charge density models have achieved transferability on molecular dynamics trajectories and diverse molecular datasets. However, kernel methods scale cubically with the size of dataset, making it unfeasible to tackle with the full materials space. In recent years message-passing neural networks—which update iteratively node and edge features of a graph by aggregate neighbor messages—have gained prominence in materials science due to their flexibility and transferability on large dataset, albeit at the cost of longer training time compared to kernel methods [21, 22]. MPNN charge density models can represent charge density with either basis expansion coefficients [23] or grid quantities [24–26]. Grid-based methods place fictitious probe nodes at grid points, which receive messages only from nearby atoms. Although much slower than basis-based methods, grid-based methods have higher accuracy and are not affected by the choice of basis. Finally, although charge density is itself an invariant quantity which transforms like scalar field, researchers have shown that it is necessary to incorporate high-order equivariant hidden layers to accurately address the complex chemical environments in molecules and solids [27].

Recent developments such as Δ -learning and physics-informed neural networks suggest that incorporating physical priors can improve data efficiency and model accuracy [28–30]. However, to the best of our knowledge, all ML charge density models reported so far learn charge density from scratch, using the total charge density as training target. Nevertheless, there are groups using atomic charge densities (SAED) as input of charge density models, demonstrating the potential of SAED to provide physical information for data driven tasks. In [31] the authors emulated the Kohn-Sham self-consistent iterations with SAED as the initial guess. They used a basis-based machine learning charge density model to predict the Kohn-Sham density map. Another group used SAED as input charge density to recover DFT charge density using image super-resolution techniques [32]. However, since image super-resolution methods operate convolutions over regular uniform grids instead of atomic graphs, they do not inherently preserve rotational invariance of charge density.

In this work, we propose using the difference charge density as the training target, referred to as Δ -SAED, which robustly improves accuracy at no additional cost. Using a state-of-the-art grid-based MPNN model, we demonstrated that Δ -SAED reduces the MAE of charge density for more than 99% of structures in the QM9 and NMC datasets, and for 90% in the MP dataset. On a Si allotropy dataset, we further showed that Δ -SAED almost always enhances the transferability to chemical accuracy for physical properties derived from non-self consistent DFT calculations, even when the accuracy of charge density itself is not improved. Therefore, Δ -SAED is more reliable when applied to unseen structures. We examined the effects of Δ -SAED on the radial distributions of the training targets and attributed its advantage to the physical priors from SAED, which reduce the complexity of the radial and angular dependencies perceived by the MPNN, alleviating its limited expressive power.

2 Results

2.1 Work Flow

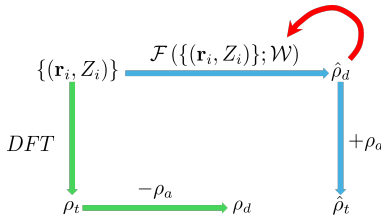


Figure 1: Workflow of Δ -SAED. Starting from a structure specified by atomic numbers and coordinates $\{(\mathbf{r}_i, Z_i)\}$, green arrows represent dataset preparation, blue arrows represent predictions using neural network $\mathcal{F}$ with parameters $\mathcal{W}$, and red arrows represent network training with the loss function $\mathcal{L}(\rho_d - \hat{\rho}_d)$.

In DFT, we are concerned with the ground-state electronic charge density $\rho_t(\mathbf{r})$. For a single isolated atom i , the charge density of its electrons $\rho_a^i(\mathbf{r} - \mathbf{r}_i)$ is localized around the nucleus $\mathbf{r}_i$. For a system of electrons and nuclei, we can define an auxiliary atomic charge density $\rho_a(\mathbf{r}) = \sum_i \rho_a^i(\mathbf{r} - \mathbf{r}_i)$ which is the superposition of isolated atomic electron densities (SAED) centered at corresponding nuclei. The difference charge density (DCD) is introduced naturally as the difference between total charge density (TCD) and SAED:

$$\rho_d(\mathbf{r}) = \rho_t(\mathbf{r}) - \rho_a(\mathbf{r}) \quad (1)$$

The Δ -SAED workflow is illustrated in Fig. 1. When preparing charge density datasets using established DFT codes, we collect not only atomic structures (model inputs) and ρ_t (ultimate goal), but also ρ_a , which can be easily extracted. During training the model parameters are optimized by minimizing $\|\hat{\rho}_d - \rho_d\|$, where $\hat{\rho}_d$ is the predicted DCD. In the inference stage, the predicted TCD $\hat{\rho}_t$ is obtained by adding ρ_a and the model-predicted $\hat{\rho}_d$.

Aiming to learn DCD defined in Eq. 1, Δ -SAED is not a trivial extension of the common trick in NNPs of training on atomization energy instead of total energy. Rather, it embodies an application of Δ -learning [28, 30], which follows the general principle that, given a physical Ansatz, the neural network only needs to learn the residual part Δ , potentially leading to improved accuracy. NNPs primarily focus on predicting forces. Since the superposition of atomic energies is independent of atomic coordinates, it does not contribute to the forces. Hence we do not consider the superposition of atomic energies as a form of physical prior in NNPs. In contrast, ρ_t has clear physical meanings, allowing us to calculate exchange-correlation potential and other related quantities. By training ρ_d as target, we effectively introduce physical prior knowledge or bias into the neural network. The analogy of SAED in NNPs is Δ -DFT [30] which uses DFT energies as physical prior of CCSD(T) energies to achieve chemical accuracies.

2.2 Performance on Benchmark Datasets

The performance of Δ -SAED is evaluated on three datasets used for benchmarking charge density models in prior studies: organic molecules (QM9) [25, 33], nickel manganese cobalt battery cathode materials (NMC) [26], and all structures with available charge density data from the Materials Project (MP) inorganic crystal database [34]. We choose Charge3Net [27], an E(3)-equivariant charge density model incorporating high-order messages, to learn the total and difference charge density. Charge3Net is grid-based, offering a more expressive representation of density compared to basis-based methods.

Following [7, 26, 27], ε_{mae} is used consistently as the evaluation metric for both TCD and DCD models, and is defined as:

$$\varepsilon_{mae} = \frac{\int_{\Omega} d^3\mathbf{r} |\rho(\mathbf{r}) - \hat{\rho}(\mathbf{r})|}{\int_{\Omega} d^3\mathbf{r} \rho_t(\mathbf{r})} \times 100\%, \quad (2)$$

where ρ and $\hat{\rho}$ denote the target and predicted total (or difference) charge densities, respectively, ρ_t denotes the target total charge density used for normalization. ε_{mae} is reported in percentage. Since the goal of ML-predicted charge density is to accelerate DFT calculations and enable further charge analyses—which require charge density information on all grid points—throughout our study we perform integrations over the full set of grid points used in the DFT calculations unless otherwise noted. For a given structure, Δ -SAED is assessed by the relative reduction of ε_{mae} :

$\Delta_{rel}\varepsilon_{mae} = \frac{\varepsilon_{mae}^{TCD} - \varepsilon_{mae}^{DCD}}{\varepsilon_{mae}^{TCD}}$, where ε_{mae}^{TCD} (ε_{mae}^{DCD}) is the density error of total (difference) charge density.

Table 1: Performance of Δ -SAED on benchmark datasets.

Dataset	#Test	Ave. ε_{mae}^{TCD}	Ave. ε_{mae}^{DCD}	Ave. $\Delta_{rel}\varepsilon_{mae}$	% Improved
NMC	500	0.056	0.046	17.0%	99.4%
QM9	10000	0.289	0.186	33.6%	99.4%
MP	2000	0.590	0.515	14.4%	89.1%

The result of DCD and TCD models in Table 1 together with the $\Delta_{rel}\varepsilon_{mae}$ distribution via violin plot in Fig 2a demonstrate the effectiveness of Δ -SAED across three benchmark datasets. The validity of the results are ensured by careful contrast settings which will be elaborated in Sec 4.2. Predicting ρ_d yields lower average ε_{mae} consistently, with an average of 17.0% $\Delta_{rel}\varepsilon_{mae}$ for the NMC dataset, 33.6% for the QM9 dataset, and 14.4% for the MP dataset. Moreover, the percentage of test structures showing improved prediction accuracy exceeds 99% for both the QM9 and NMC datasets, and remains as high as 89.1% for the more diverse MP dataset, further demonstrating the capability of Δ -SAED as a simple yet robust method to enhance accuracy over baseline approaches. Notably, there are a few outliers for the NMC and the MP dataset, which show much less accuracy of ε_{mae} after imposing Δ -SAED technique. We hypothesize that this is due to the rarity of the corresponding chemical environments in the training data, which may be mitigated by employing training strategies such as active learning [35].

2.3 Improved Transferability for Properties from Non-Self Consistent Calculations

Using a Si allotropy dataset composed of diverse crystal structures from Materials Project [36], we report the improved transferability of Δ -SAED by comparing the MAEs of per-atom energies, forces, band gaps and band energies derived from non-self-consistent calculations with predicted charge densities. We first collected 42 Si allotropes from Materials

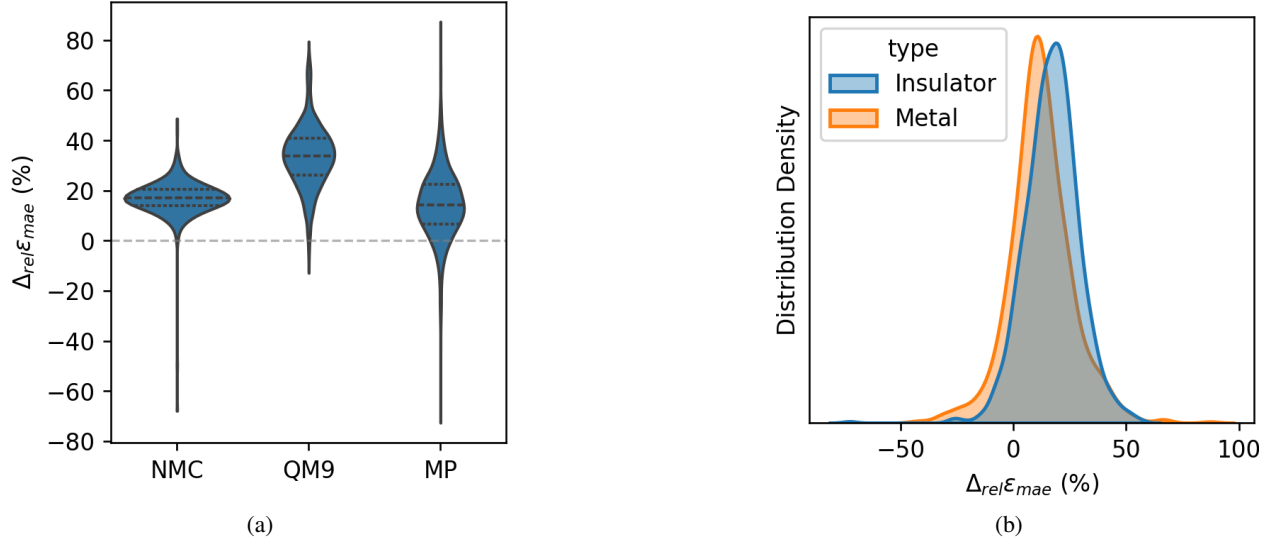


Figure 2: Δ -SAED performance on benchmark datasets with $\Delta_{rel}\epsilon_{mae}$ distributions across test structures. (a) Distributions of $\Delta_{rel}\epsilon_{mae}$ via violin plots for the NMC, QM9, MP test set. (b) $\Delta_{rel}\epsilon_{mae}$ distribution for metals and insulators in the MP test set.

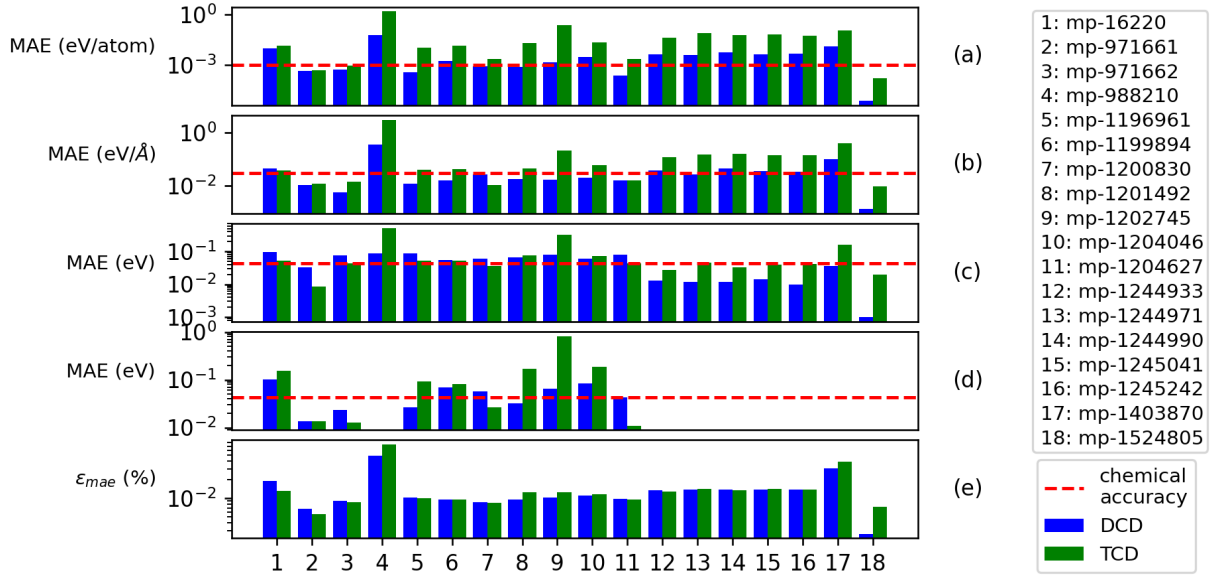


Figure 3: MAEs of per-atom energies (a), forces (b), band gaps (c), band energies (d) and charge density (e) for 18 Si allotropes from Materials Project to evaluate transferability of the Si DCD and TCD models. The MP ids are shown in the right box. Blue bars correspond to non-self-consistent calculations derived from $\rho_d + \rho_a$, while green bars from ρ_t . All MAEs use self-consistent results as the ground truth. The red lines are chemical accuracies, with 1 meV/atom for per-atom energy, 30 meV/Å for force, and 43 meV for band energy and band gap.

Project, 24 of which were used to construct the Si dataset, while the remaining 18 allotropes were reserved for transferability evaluation. The 24 Si allotropes for the dataset, among which one-fourth are insulators, contain no more than 60 atoms and exhibit good stability. Twelve of these were further expanded into larger supercells. We then introduced random displacements of 0.1 Å to the original structures 30 times, generating a total of 1080 structures. The Si dataset was then split into training, validation, and test sets with a ratio of 8:1:1.

Using identical hyperparameters, the DCD (TCD) models yield MAEs of 0.02 meV/atom (0.2 meV/atom), 0.5 meV/Å (1.5 meV/Å), 0.4 meV (0.9 meV), 0.06 meV (0.1 meV), for per-atom energy, force, band energy and band gap respectively, while the commonly adopted chemical accuracy thresholds are 1 meV/atom, 30 meV/Å, 42 meV and 43 meV. Not surprisingly, both the DCD and TCD models achieved chemical accuracy predicting random displaced structures in the training set, with the DCD model yielding slightly lower MAEs, although the average density error ε_{mae} is similar (0.3%) for both models. This suggests the comparable capability and even potential advantages of Δ -SAED in related applications such as phonon calculations, where for ML models good transferability is required for similar chemical environment with larger system size.

We further evaluate the models’ transferability with 18 Si allotropes in MP outside of the training set. The average ε_{mae} of DCD model (1.42%) is slightly smaller than that of TCD model (1.62%), which is consistent with the trend in Table 1. The MAEs of non-self-consistent properties are presented in Fig 3. For mechanical properties, the average per-atom energy MAEs are 6.4 meV/atom (DCD) and 120 meV/atom (TCD). The average force MAEs are 46 meV/Å (DCD) and 260 meV/Å (TCD). The force MAEs of the DCD model is close to the common threshold for structure relaxations (30 meV/Å). For electronic properties, the average band energy MAEs are 48 meV (DCD) and 92 meV (TCD), while the band gap MAEs are 29 meV (DCD) and 87 meV (TCD). The band gap MAEs for some structures are exactly zero, indicating that their metallic states are predicted correctly for both DCD and TCD models. Our results indicate that models trained on DCD exhibit superior transferability over which trained on TCD, suggesting DCD models are more reliable when confronting with unseen structures.

2.4 DCD Exhibits Simpler Radial and Angular Dependencies in the Core Region

In this section, we investigate why and to what extent Δ -SAED is beneficial. We attribute its advantage to the simpler radial and angular dependencies of DCD in the core region. In Fig. 2b the distribution of the $\Delta_{rel}\varepsilon_{mae}$ for metals and insulators in the MP dataset, which occur in approximately equal proportions, are present separately. Δ -SAED achieves better performance on insulators, with an average $\Delta_{rel}\varepsilon_{mae}$ of 17.2%, compared to 11.3% for metals. The performance disparity of Δ -SAED among insulators and metals suggests that Δ -SAED works in line with the Δ -learning principle, i.e., incorporating physical priors would smooth the regression task. This is evidenced by the empirical observation that insulators’ ρ_d is closer to zero than metals’ due to less complex charge transfer process, indicating that for insulators the smoothness advantage of ρ_d over ρ_t is more prominent than that of metals.

In Fig. 4 we displays 4 Si allotropes’ radial distributions, defined by grid points within 2 Å of each atom, for ρ_t , ρ_d and their corresponding absolute error $|\rho_t - \hat{\rho}_t|$, $|\rho_d - \hat{\rho}_d|$. We choose mp-1202745 and mp-988210 as representatives of the 18 test structures in Fig. 3 because their MAEs are close to chemical accuracy and their radial distributions resemble those of other test structures, while mp-988210 and mp-1403870 are chosen as examples with outlier errors. Both TCD and DCD models struggle to predict the charge density and properties for the two allotropes, because the MPNN charge density model has less transferability for complex geometry chemical environment than complex compound combinations, as also reported in [37], though Δ -SAED still outperforms. Overall, ρ_d is smoother as function of r and shows a narrower spread, indicating simpler angular and radial dependencies. This confirms that SAED provides useful physical prior information, making ρ_d easier to learn. The radial error distributions in Fig. 4a, 4b show that, near atomic centers, the error increases for the TCD model but decreases for the DCD model. As expected, Δ -SAED performs better around atomic centers, although the overall average error of DCD is not always lower than that of TCD (mp-1244933 for example). Remarkably, even in such cases, the DCD model achieves better performance on structural properties. This suggests Δ -SAED’s reliability for non-self-consistent calculations, and the limitation of relying solely on the average charge density error as an evaluation metric for charge density models.

In Fig. 5 we show the per-atom ε_{mae} of the MP test set, by integrating within a sphere of radius 2 Å centered at each atom using Eq. 1. We find Δ -SAED improves across the periodic table consistently, with an average 0.08% of absolute improvement $\Delta\varepsilon_{mae} = \varepsilon_{mae}^{TCD} - \varepsilon_{mae}^{DCD}$, despite as element gets heavier the sampling probability declines. The improvement is not because SAED compensates for inefficient uniform sampling of the TCD model around the core, as the radial error distributions for Si allotrope in Fig. 4 shows that TCD model can predict core charge densities reasonably well, though not as accurately as DCD model does. Rather, the improvement reflects the limited expressive power of the model itself in capturing the complex radial and angular dependence of ρ_t , and SAED provides useful physical priors across the periodic table to compensate for the model architecture’s expressive limitations.

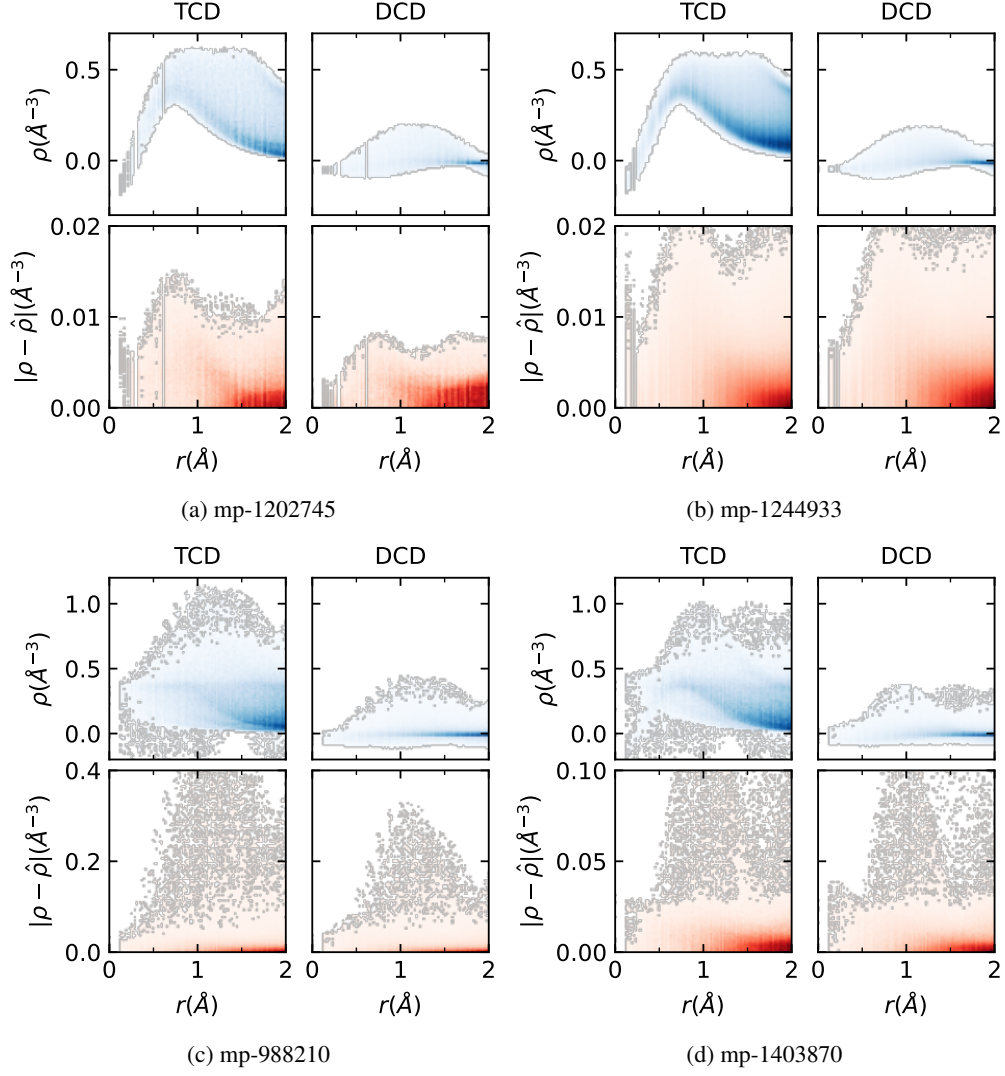


Figure 4: Radial distributions of density functions for 4 Si allotropes in Fig. 3. Blue denotes the target density, red denotes the absolute density error, and darker regions indicate higher grid density. TCD refers to total charge density, with ρ_t shown in blue and $|\rho_t - \hat{\rho}_t|$ in red, while DCD refers to difference charge density, with ρ_d shown in blue and $|\rho_d - \hat{\rho}_d|$ in red.

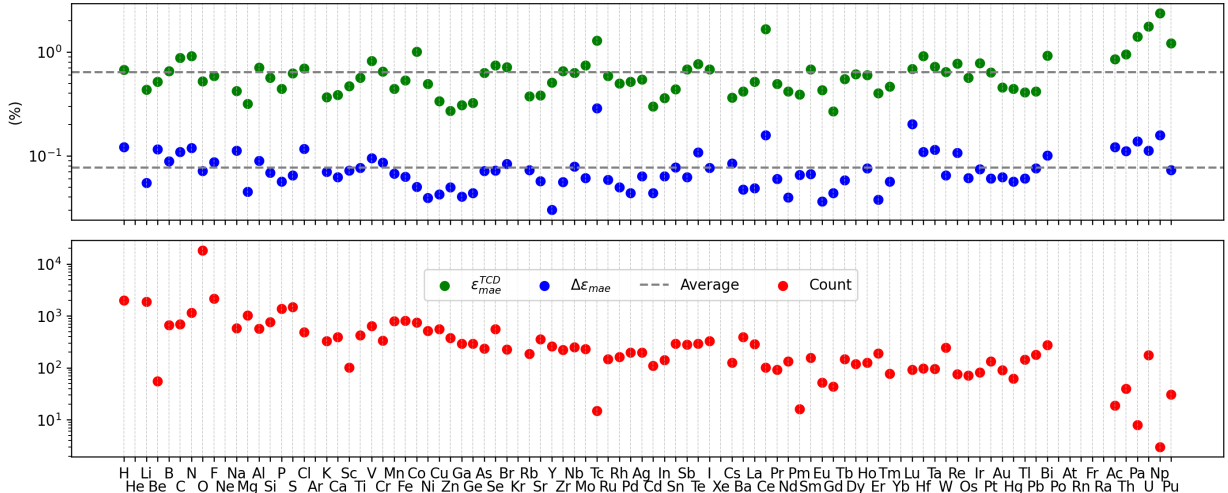


Figure 5: Per-atom ε_{mae} and the count of each element type of the MP test set.

3 Discussion

This work introduces Δ -SAED, a robust and cost-free method to improve the accuracy and transferability of charge density and derived properties beyond baseline models. Inspired by the spirit of Δ -learning, Δ -SAED applies superposition of atomic electron density as physical priors and trains on the difference charge density instead of the total charge density. This reduces the complexity of training target’s radial and angular dependencies at core regions, mitigating machine learning models’ limited expressiveness in complex fitting task. Although advantages have so far been demonstrated only on grid-based architectures, we expect Δ -SAED to also work for basis-based architectures, as it provides an exact component of the charge density, bypassing the less accurate basis-expansion formulation.

Looking forward, Δ -SAED paves the way for building charge density foundation models that integrates calculations from different DFT codes. Given a structure and functional, TCD is not unique in different DFT implementations due to approximations like pseudopotentials and basis sets. In contrast, DCD is, in principle, free from this issue once calculations converge, since it captures the charge transfer from SAED—a well defined state by solving the radial Schrödinger equations for atoms individually.

4 Methods

4.1 Datasets

The total charge densities of the NMC and QM9 datasets are provided by Jørgensen and Bhowmik [26, 30]. The total charge densities of the MP dataset are calculated by [34]. The NMC dataset consists of 2000 randomly sampled nickel manganese cobalt oxides containing varying levels of lithium content, with training, validation, and test splits of size 1450, 50, and 500 respectively. The QM9 dataset contains 133845 small organic molecules, with training, validation, and test splits of size 123835, 50, and 10,000 respectively. The MP dataset contains 117173 structures, splited randomly into training, validation, and test sets with sizes 114661, 512, and 2000 respectively. Our MP dataset is different from the one used in [27] because structures with available charge density information are updated frequently.

4.2 Training

The training settings, including model hyper parameters and optimizing schedules, follow those reported by Koker [27]. For every dataset, the settings used for training ρ_d are the corresponding same settings for ρ_t . The training setting for the Si dataset is identical to that of the NMC dataset. The model parameters for the NMC TCD and QM9 TCD models are adopted from [27]. Notably, the average ε_{mae} for TCD models in Tab. 1 differs from those reported by Koker, as our evaluation uses all probing points in line with DFT, whereas Koker employed a random sampling for the integral in Eq. 2. The structures for the Si allotropy dataset and all trained model parameters are available on Zenodo (10.5281/zenodo.17180513).

4.3 DFT Calculations

All DFT calculations were performed using plane wave based VASP code [38] with PBE functional [39] and projector-augmented wave (PAW) pseudopotentials [40, 41]. All Si allotropes are calculated with a $6 \times 6 \times 6$ Monkhorst-Pack k-point mesh [42]. After performing self-consistent-field calculations, ρ_a is computed using the same pseudopotential, with the electronic relaxation disabled (NELM = 0), the charge density initialization scheme set to SAED(ICHARG = 12), the DFT grid forced consistent. Notably, each element in the MP dataset uses a unique, consistent pseudopotential across all structures, avoiding any physical discrepancies due to pseudopotential mixing.

Acknowledgements

We acknowledge financial support from the National Key R&D Program of China (No. 2022YFA1402901), NSFC (grants No. 12188101), Shanghai Science and Technology Program (No. 23JC1400900), the Guangdong Major Project of the Basic and Applied Basic Research (Future functional materials under extreme conditions–2021B0301030005), Shanghai Pilot Program for Basic Research—Fudan University 21TQ1400100 (23TQ017), the robotic AI-Scientist platform of Chinese Academy of Science, and New Cornerstone Science Foundation.

References

- [1] Keith T. Butler, Daniel W. Davies, Hugh Cartwright, Olexandr Isayev, and Aron Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547–555, July 2018.
- [2] W. Kohn. Density functional and density matrix method scaling linearly with the number of atoms. *Physical Review Letters*, 76(17):3168–3171, April 1996.
- [3] Kurt Lejaeghere, Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E. Castelli, Stewart J. Clark, Andrea Dal Corso, Stefano de Gironcoli, Thierry Deutsch, John Kay Dewhurst, Igor Di Marco, Claudia Draxl, Marcin Dułak, Olle Eriksson, José A. Flores-Livas, Kevin F. Garrity, Luigi Genovese, Paolo Giannozzi, Matteo Giantomassi, Stefan Goedecker, Xavier Gonze, Oscar Grånäs, E. K. U. Gross, Andris Gulans, François Gygi, D. R. Hamann, Phil J. Hasnip, N. A. W. Holzwarth, Diana Iuşan, Dominik B. Jochym, François Jollet, Daniel Jones, Georg Kresse, Klaus Koepernik, Emine Küçükbenli, Yaroslav O. Kvashnin, Inka L. M. Locht, Sven Lubeck, Martijn Marsman, Nicola Marzari, Ulrike Nitzsche, Lars Nordström, Taisuke Ozaki, Lorenzo Paulatto, Chris J. Pickard, Ward Poelmans, Matt I. J. Probert, Keith Refson, Manuel Richter, Gian-Marco Rignanese, Santanu Saha, Matthias Scheffler, Martin Schlipf, Karlheinz Schwarz, Sangeeta Sharma, Francesca Tavazza, Patrik Thunström, Alexandre Tkatchenko, Marc Torrent, David Vanderbilt, Michiel J. van Setten, Veronique Van Speybroeck, John M. Wills, Jonathan R. Yates, Guo-Xu Zhang, and Stefaan Cottenier. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280), March 2016.
- [4] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 140(4A):A1133–A1138, November 1965.
- [5] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 98(14), April 2007.
- [6] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Physical Review Letters*, 104(13), April 2010.
- [7] Felix Brockherde, Leslie Vogt, Li Li, Mark E. Tuckerman, Kieron Burke, and Klaus-Robert Müller. Bypassing the kohn-sham equations with machine learning. *Nature Communications*, 8(1), October 2017.
- [8] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136(3B):B864–B871, November 1964.
- [9] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133–A1138, November 1965.
- [10] Beatriz G. Del Rio, Brandon Phan, and Rampi Ramprasad. A deep learning framework to emulate density functional theory. *npj Computational Materials*, 9(1):158, August 2023.
- [11] Masashi Tsubaki and Teruyasu Mizoguchi. Quantum deep field: Data-driven wave function, electron density generation, and atomization energy prediction and extrapolation with machine learning. *Physical Review Letters*, 125(20), November 2020.
- [12] Graeme Henkelman, Andri Arnaldsson, and Hannes Jónsson. A fast and robust algorithm for bader decomposition of charge density. *Computational Materials Science*, 36(3):354–360, June 2006.

-
- [13] Andreas Savin, Reinhard Nesper, Steffen Wengert, and Thomas F. Fässler. Elf: The electron localization function. *Angewandte Chemie International Edition in English*, 36(17):1808–1832, September 1997.
- [14] Ronald E. Cohen. Origin of ferroelectricity in perovskite oxides. *Nature*, 358(6382):136–138, July 1992.
- [15] Heleen van Gog and Marijn A. van Huis. Structural and electronic properties of frenkel and schottky defects at the mgo100 surface: Spin polarization, mid-band gap states, and charge trapping at vacancy sites. *The Journal of Physical Chemistry C*, 123(23):14408–14420, May 2019.
- [16] Veaceslav Coropceanu, Xian-Kai Chen, Tonghui Wang, Zilong Zheng, and Jean-Luc Brédas. Charge-transfer electronic states in organic solar cells. *Nature Reviews Materials*, 4(11):689–707, September 2019.
- [17] Aleksandr O. Kondrakov, Holger Geßwein, Kristina Galdina, Lea de Biasi, Velimir Meded, Elena O. Filatova, Gerhard Schumacher, Wolfgang Wenzel, Pascal Hartmann, Torsten Brezesinski, and Jürgen Janek. Charge-transfer-induced lattice collapse in ni-rich ncm cathode materials during delithiation. *The Journal of Physical Chemistry C*, 121(44):24381–24388, October 2017.
- [18] Andrea Grisafi, Alberto Fabrizio, Benjamin Meyer, David M. Wilkins, Clemence Corminboeuf, and Michele Ceriotti. Transferable machine-learning model of the electron density. *ACS Central Science*, 5(1):57–64, December 2018.
- [19] Alberto Fabrizio, Andrea Grisafi, Benjamin Meyer, Michele Ceriotti, and Clemence Corminboeuf. Electron density learning of non-covalent systems. *Chemical Science*, 10(41):9424–9432, 2019.
- [20] Bruno Focassio, Michelangelo Domina, Urvesh Patil, Adalberto Fazzio, and Stefano Sanvito. Linear Jacobi-Legendre expansion of the charge density for machine learning-accelerated electronic structure calculations. *npj Computational Materials*, 9(1):87, May 2023.
- [21] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry, 2017.
- [22] Tian Xie and Jeffrey C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(14), April 2018.
- [23] Chaoqiang Feng, Yaolong Zhang, and Bin Jiang. Efficient Sampling for Machine Learning Electron Density and Its Response in Real Space. *Journal of Chemical Theory and Computation*, 21(2):691–702, January 2025.
- [24] Sheng Gong, Tian Xie, Taishan Zhu, Shuo Wang, Eric R. Fadel, Yawei Li, and Jeffrey C. Grossman. Predicting charge density distribution of materials using a local-environment-based graph convolutional network. *Physical Review B*, 100(18), November 2019.
- [25] Peter Bjørn Jørgensen and Arghya Bhowmik. DeepDFT: Neural Message Passing Network for Accurate Charge Density Prediction, 2020.
- [26] Peter Bjørn Jørgensen and Arghya Bhowmik. Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids. *npj Computational Materials*, 8(1):183, August 2022.
- [27] Teddy Koker, Keegan Quigley, Eric Taw, Kevin Tibbetts, and Lin Li. Higher-order equivariant neural networks for charge density prediction in materials. *npj Computational Materials*, 10(1):161, July 2024.
- [28] Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole von Lilienfeld. Big data meets quantum chemistry approximations: The δ -machine learning approach. *Journal of Chemical Theory and Computation*, 11(5):2087–2096, April 2015.
- [29] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, February 2019.
- [30] Mihail Bogojeski, Leslie Vogt-Maranto, Mark E. Tuckerman, Klaus-Robert Müller, and Kieron Burke. Quantum chemical accuracy from density functional approximations via machine learning. *Nature Communications*, 11(1), October 2020.
- [31] Feitong Song and Ji Feng. NeuralSCF: Neural network self-consistent fields for density functional theory, June 2024.
- [32] Chenghan Li, Or Sharir, Shunyue Yuan, and Garnet Kin-Lic Chan. Image super-resolution inspired electron density prediction. *Nature Communications*, 16(1), May 2025.
- [33] Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole Von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1(1):140022, August 2014.

-
- [34] Jimmy-Xuan Shen, Jason M. Munro, Matthew K. Horton, Patrick Huck, Shyam Dwaraknath, and Kristin A. Persson. A representation-independent electronic charge density database for crystalline materials. *Scientific Data*, 9(1):661, October 2022.
- [35] Linfeng Zhang, De-Ye Lin, Han Wang, Roberto Car, and Weinan E. Active learning of uniformly accurate interatomic potentials for materials simulation. *Physical Review Materials*, 3(2):023804, February 2019.
- [36] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin a. Persson. The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 2013.
- [37] Qin Xuejian, Lv Taoyuze, and Zhong Zhicheng. EAC-Net: Real-space charge density via equivariant atomic contributions, August 2025.
- [38] G. Kresse and J. Furthmüller. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, October 1996.
- [39] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, October 1996.
- [40] P. E. Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953–17979, December 1994.
- [41] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B*, 59(3):1758–1775, January 1999.
- [42] Hendrik J. Monkhorst and James D. Pack. Special points for Brillouin-zone integrations. *Physical Review B*, 13(12):5188–5192, June 1976.